Master's Thesis

Unsupervised Training Of Denoisers For Low-Dose CT Reconstruction Without Full-Dose Ground Truth

Kwanyoung Kim

Department of Electrical Engineering

Graduate School of UNIST

2020

Unsupervised Training Of Denoisers For Low-Dose CT Reconstruction Without Full-Dose Ground Truth

Kwanyoung Kim

Department of Electrical Engineering

Graduate School of UNIST

Unsupervised Training Of Denoisers For Low-Dose CT Reconstruction Without Full-Dose Ground Truth

thesis/dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

Kwanyoung Kim

June 15, 2020
Approved by

Advisor
Se Young Chun

Unsupervised Training Of Denoisers For Low-Dose CT Reconstruction Without Full-Dose Ground Truth

Kwanyoung Kim

This certifies that the thesis/dissertation of Kwanyoung Kim is approved.

June 15, 2020

signatrue

Advisor : Se Young Chun

signatrue

Seungjoon Yang : Thesis Committee Member #1

signatrue

Jeong hwan Jeon : Thesis Committee Member #2

Abstract

Recently, deep neural network (DNN) based methods for low-dose CT have been investigated to achieve excellent performance in both image quality and computational speed. However, almost all methods using DNNs for low-dose CT require clean ground truth data with full radiation dose to train the DNNs. In this work, we attempt to train DNNs for low-dose CT reconstructions with reduced tube current by investigating unsupervised training of DNNs for denoising sensor measurements or sinograms without full-dose ground truth images. In other words, our proposed methods allow training of DNNs with only noisy low-dose CT measurements. First, the Poisson Unbiased Risk Estimator (PURE) is investigated to train a DNN for denoising CT measurements, and a method is proposed for reconstructing the CT image using filtered back-projection (FBP) and the DNN trained with PURE. Then, the CT forward model-based Weighted Stein's Unbiased Risk Estimator (WSURE) is proposed to train a DNN for denoising CT sinograms and to subsequently reconstruct the CT image using FBP. Our proposed methods achieve excellent performance in both fast computation and reconstructed image quality, which is more comparable to the results of the DNNs trained with full-dose ground truth data than other state-of-the-art denoising methods such as the BM3D, Deep Image Prior, and Deep Decoder.

Contents

List of Figures

List of Tables

CHAPTER I

Introduction

1.1 Computed Tomography

In this section, we briefly introduce the theorerical background of computed tomography such as geometry, Randon trasnform in 2D, Fourier-slice theorem, Filter-Back Projection algorithm.

1.1.1 CT Geometry

In general, the geometry of Computed Tomography(CT) can be roughly divided into fan beam, cone beam geometry. The number of detector's row of fan beam is single so that the shape of detector looks like line. However, cone beam have multiple detector's row and the shape of detecor is flat panel as shown in Fig 1.1.

1.1.2 Radon transform in 2D

To obtain projection data from the target object $f(x, y)$ in 2D or 3D geometry, the Radon transform method that relates $f(x, y)$ with the collection of line intgral is used in computed tomography. Fig. 1.2 shows the geometry of the line integrals related with the 2D Radon transform. In 2D geometry, we can define the line $\mathcal{L}(r, \varphi)$ in the Euclidean plane as following:

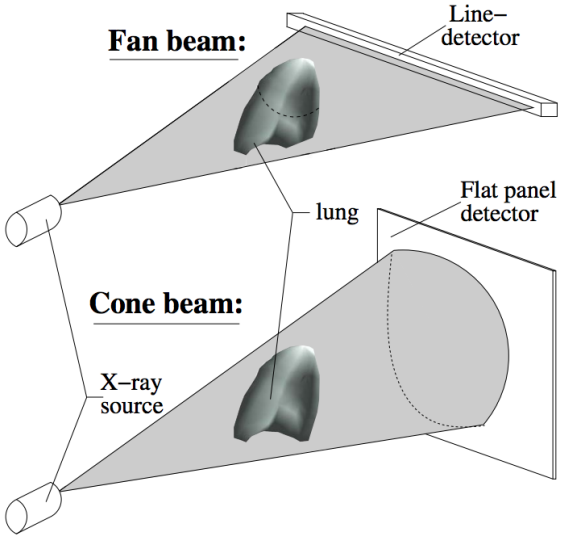


Figure 1.1: Visualization of difference between fan beam and cone beam.

$$\mathcal{L}(r,\varphi) = \{(x,y) \in \mathbb{R}^2 : x\cos\varphi + y\sin\varphi = r\} \tag{I.1}$$
$$= \{(x,y) \in \mathbb{R}^2 : (x,y)\cdot(\cos\varphi,\sin\varphi) = r\} \tag{I.2}$$
$$= \{(r\cos\varphi - l\sin\varphi, r\sin\varphi + l\cos\varphi) : l \in \mathbb{R}\} \tag{I.3}$$

where φ is angle counter-clockwise from the y axis and at a signed distrance r from the origin. The projection data $p_\varphi(r)$ can be defined as the ilne integral through $f(x,y)$ along the line $\mathcal{L}(r,\varphi)$ and can be described as following:

$$p_\varphi(r) = \int_{\mathcal{L}(r,\varphi)} f(x,y)\, dl \tag{I.4}$$
$$= \int_{-\infty}^{\infty} f(r\cos\varphi - l\sin\varphi, r\sin\varphi + l\cos\varphi)\, dl \tag{I.5}$$
$$= \iint_{-\infty}^{\infty} f(x,y)\delta(x\cos\varphi + y\sin\varphi - r)\, dxdy \tag{I.6}$$

where $\delta(\cdot)$ denotes the 1D Dirac impulse. In other words, the projection data $p_\varphi(r)$ can be expressed as Radon transfrom of $f(x,y)$. The projection data $p_\varphi(r)$ is function of two variable r,φ so that this data can be displayed 2D grayscale image, so called sinogram. In general, the purpose of computed tomography reconstruction method is to predict the object $f(x,y)$ given a measured sinogram $p_\varphi(r)$.

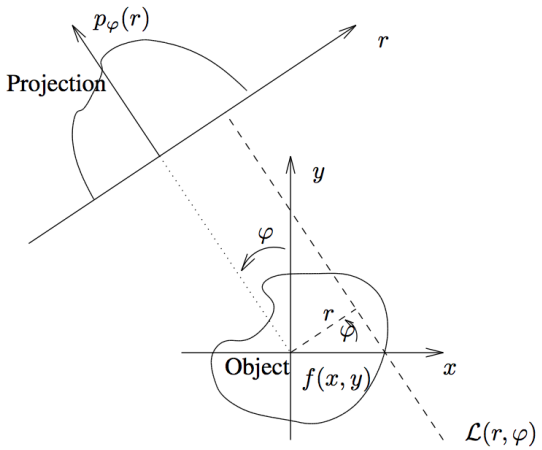


Figure 1.2: Explanation of the line integral related to the Radon transform [1].

1.1.3 Fourier-slice theorem

To analyze projection data $p_\varphi(r)$ in Fourier domain, Fourier-slice theorem which 1D Fourier transform of $p_\varphi(r)$ is equal to the 2D Fourier transfrom of $f(x,y)$ at angle φ is used.
Firstly, the 1D Fourier transform of $p_\varphi(r)$ can be expressed as :

$$P_\varphi(v) = \int_{-\infty}^{\infty} p_\varphi(r) \exp^{-j2\pi vr} \; dr \tag{I.7}$$

The 2D Fourier transform of $f(x,y)$ can be described as :

$$F(u,v) = \iint_{-\infty}^{\infty} f(x,y) \exp^{-j2\pi(xu+vy)} \; dxdy \tag{I.8}$$

Finally, we can simply represent Fourier-slice theorem :

$$P_\varphi(v) = F(v\cos\varphi, v\sin\varphi) \tag{I.9}$$

1.1.4 Filter-Back Projection (FBP)

The purpose of computed tomographic image reconstruction is recover the object $f(x,y)$ from $p_\varphi(r)$ that is obtained by Radon transform, which is so called backprojection. When $f(x,y)$ is noiseless case, the direct Fourier reconstruction method can be applied to recover the target object $f(x,y)$. However, when the object $f(x,y)$ is contaminated by noise or blurred, the method that reduce noise or blurr is required. Among these apporaches, we introduce Filter-

Back Projection method that is widely used in CT field. FBP method that apply Ramp filters to projection data $p_\varphi(r)$ in frequency domain and use back projection method. Because of this sequnece, this method is called Filter-Back Projection and the pipeline of FBP method can be represented as :

$$f(x,y) \rightarrow \boxed{\text{Projection}} \rightarrow p_\varphi(r) \rightarrow \boxed{\text{Ramp filters}} \rightarrow \check{p}_\varphi(r) \rightarrow \boxed{\text{Backprojection}} \rightarrow f(x,y)\,.$$

By using the Fourier-slice theorem, Filter-Back Projection method can be described as follows:

$$f(x,y) = \iint F(u,v)\exp^{i2\pi(xu+yv)}\ dudv \tag{I.10}$$

$$= \int_0^\pi \int_{-\infty}^{\infty} F(v\cos\varphi, v\sin\varphi)\exp^{i2\pi v(x\cos\varphi+y\sin\varphi)}\,|v|\ dvd\varphi \tag{I.11}$$

$$= \int_0^\pi \int_{-\infty}^{\infty} P_\varphi(v)\exp^{i2\pi v(x\cos\varphi+y\sin\varphi)}\,|v|\ dvd\varphi \tag{I.12}$$

$$= \int_0^\pi \hat{p}_\varphi(x\cos\varphi + y\sin\varphi)\ d\varphi \tag{I.13}$$

where the filtered projection data $\hat{p}_\varphi(r)$ is defined as follows:

$$\hat{p}_\varphi(r) = \int_{-\infty}^{\infty} P_\varphi(v)\exp^{i2\pi vr}\ dv \tag{I.14}$$

1.2 Introduction

Low-dose CT has been widely investigated in response to the concerns on patient safety in full-dose CT scans. Typically, low-dose CT has been implemented either by reducing the number of projection views (sparse view CT) based on compressed sampling theories or by reducing the X-ray tube current / voltage using statistical image reconstructions with accurate noise models. However, one of the drawbacks of conventional low-dose CT works is the relatively slow reconstruction speed of the iterative optimizations, even with state-of-the-art algorithms [4–8]. This is a critical drawback for emergency CT scanning where real-time image reconstruction is essential for patient care. Thus, low-dose CT with iterative reconstruction has not been successful in replacing full-dose CT with fast filtered back-projection (FBP) in clinics yet.

Recently, deep neural network (DNN) based reconstruction methods for low-dose CT have significantly improved the image quality and reconstruction speed, enabling nearly real-time

reconstruction. There are three groups of approaches for low-dose CT reconstructions, namely in the image, measurement (sensor raw data, count), and sinogram domains.

The image domain approach for low-dose CT reconstructions is the most popular. One approach involves the reconstruction of a contaminated CT image from sparse-view sinograms first using FBP, and then a DNN is used to remove streaking artifacts to yield the final CT image [9–12]. Another approach involves the reconstruction of a noisy CT image from noisy sinograms, and then a DNN is used to reduce colored noise to produce the final CT image [13–17]. Lastly, there are hybrid methods using iterative algorithms with learned priors to ensure sinogram data consistency [18–20]. Note that model based DNN image recovery methods have been investigated in magnetic resonance imaging (MRI) [21–23], which can be potentially useful for low-dose CT reconstruction. In addition, methods incorporating various loss functions such as structual loss and adversarial loss, have been proposed to enhance the image quality of low-dose CT reconstruction [24–27].

Noise reduction approaches in the measurement / sinogram domains for low-dose CT have also been investigated such as DNN-based filtering of CT measurements [28] and DNN-based noise reduction methods in the sinogram domain [29–31]. Direct reconstruction from low-dose CT sinograms has also been investigated [32, 33].

One of the common requirements of the aforementioned methods to train a DNN for low-dose CT is the availability of clean full-dose ground truth data. Currently, many full-dose CT datasets are available for CT reconstruction such as the 2016 Low-Dose CT Grand Challenge dataset [34]. However, we argue that more CT datasets are required for different scanners, applications, and improved performance. Due to the advanced CT scanner technology (e.g., detector resolution), most current DNN based low-dose CT reconstruction methods require newly collected full-dose CT datasets to achieve better image quality. The acquisition of more full-dose data might also be required if new image features from new diseases should be incorporated into the DNN-based low-dose CT reconstruction methods. While the collection of new full-dose CT images is feasible, this is undesirable especially in the case of high-risk patients (*e.g.*, children). Unsupervised training methods without full-dose ground truth datasets may provide an alternative "safe" option for these challenges via the acquisition of low-dose CT datasets.

Over the last couple of years, various works / studies have investigated self-supervised / unsupervised training methods for DNNs. The deep image prior (DIP) exploits the structure of a DNN to regularize a given noisy image for each test image [35]. This work has been incorporated into positron emission tomography (PET) image reconstruction [36]. The DIP does not train a network in advance, but uses an iterative algorithm with the network to yield a denoised image. Thus, the DIP is slow compared with other DNN methods. To avoid the issue

of overfitting in the DIP, DeepRED was proposed to improve the DIP by adding regularization of denoising (RED) [37], while the Deep Decoder was investigated as another enhanced version of the DIP, which have an under-parametrized neural network with no convolution [38].

Noise2Noise was proposed to train DNNs for image denoising with the two independent noise realizations per image in an unsupervised method [39]. There have been several studies unsupervised training with only one noise realization per image. Stein's unbiased risk estimator (SURE) based training method for Gaussian noise was proposed to train DNNs with a single noise realization per image [40, 41]. This work has been extended to train DNNs with a set of undersampled compressive sensing (CS) measurements [42]. Although Noise2Void was proposed for training denoiser DNNs using a blind-spot network and yielded good performance, it did not provide better performance than conventional methods such as BM3D [43], especially for low noise levels [44]. Noise2Self was also proposed to train denoiser DNNs using J-invariant masks. However, optimal selection of the J-partition is challenging [45]. Noise2Void was further improved, yielding comparable performance to the network trained with ground truth data for Gaussian denoising [46].

In this work, we investigate unsupervised training of DNN denoisers for low-dose CT reconstructions without using full-dose ground truth CT data. First, we investigate supervised DNN based low-dose CT reconstruction frameworks by denoising in the image, sinogram and measurement domains as oracle cases for our unsupervised training to aid comparison. Then, employing the Monte-Carlo Poisson Unbiased Risk Estimator (MC-PURE) expression [47], we propose unsupervised training methods for low-dose CT in the measurement domain. Furthermore, extending our previous work on unsupervised training of denoisers for Gaussian noise using SURE [40], we propose unsupervised training methods for low-dose CT using our proposed Weighted Stein's Unbiased Risk Estimator (WSURE) with a model-based weighted Gaussian approximation for sinograms [48]. Note that denoised CT measurements / sinograms allow FBP to use the ramp filter for passing more high frequency information than before attenuating noise. Extensive experiments were performed on the AAPM CT image dataset [34] and our proposed methods were compared with other state-of-the-art methods such as the BM3D [43], DIP [35], and Deep Decoder [38].

CHAPTER II

Background

2.1 Background

2.1.1 Supervised training of DNN-based denoisers

Assume that the ground truth signal $\boldsymbol{x}$ is contaminated by an additive white Gaussian noise $\boldsymbol{n} \in \mathbb{R}^K$:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}, \qquad \boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{II.1}$$

where $\boldsymbol{x} \in \mathbb{R}^K$ is the unknown signal, $\boldsymbol{y} \in \mathbb{R}^K$ is the noisy observation, σ^2 is the known variance and $\boldsymbol{I}$ is an identity matrix. Let $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ be denoted as $\boldsymbol{n} \sim \mathcal{N}_{0,\sigma^2}$. An estimator of $\boldsymbol{x}$ from $\boldsymbol{y}$, or a denoiser, can be described as a function $\boldsymbol{f}(\boldsymbol{y})$ of $\boldsymbol{y}$ where $\boldsymbol{f}$ is a function from $\mathbb{R}^K$ to $\mathbb{R}^K$.

DNN-based denoisers, which have demonstrated excellent performances over conventional methods [49, 50], can be typically modeled as a parameterized function $\boldsymbol{f}(\boldsymbol{y}; \boldsymbol{\theta})$ where $\boldsymbol{y}$, generated based on (II.1), is an input vector to the network and $\boldsymbol{\theta} \in \mathbb{R}^P$ is a weight vector of the DNN. The weights $\boldsymbol{\theta}$ are determined through training by minimizing the loss function.

A general empirical risk function for denoisers with the observed signal equation in (II.1) and the ground truth images $\boldsymbol{x}$ is given by:

$$\mathbb{E}_{\boldsymbol{x} \sim p(\boldsymbol{x}), \boldsymbol{n} \sim \mathcal{N}_{0,\sigma^2}} \|\boldsymbol{x} - \boldsymbol{f}(\boldsymbol{y}; \boldsymbol{\theta})\|^2. \tag{II.2}$$

With a large M, (II.2) is well-approximated by the strong law of large numbers (SLLN) using the following empirical risk:

$$\frac{1}{M}\sum_{j=1}^{M}\|\boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})-\boldsymbol{x}^{(j)}\|^2 \tag{II.3}$$

where $\{(\boldsymbol{x}^{(1)},\boldsymbol{y}^{(1)}),\cdots,(\boldsymbol{x}^{(M)},\boldsymbol{y}^{(M)})\}$ are M pairs of a training dataset, sampled from the joint distribution of $\boldsymbol{x}^{(j)}\sim p(\boldsymbol{x})$ and $\boldsymbol{n}^{(j)}\sim\mathcal{N}_{0,\sigma^2}$ using the signal generation model (II.1). Note that (II.3) is an unbiased estimator of (II.2). In practice, calculating the gradient of (II.3) for a large M is inefficient, as this gradient can be well-approximated by a small amount / number of well-shuffled training data. Thus, a mini-batch is typically used for efficient DNN training by computing the mini-batch objective function where M denotes the size of one mini-batch.

Given the ground truth images $\boldsymbol{x}^{(j)}$, supervised training of the DNN $\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})$ can be performed by minimizing (II.3) with respect to $\boldsymbol{\theta}$ using a gradient-based optimization algorithm, such as the stochastic gradient descent (SGD) [51], Nesterov momentum [52], or the Adam optimizer [53] to name a few.

Recently, in deep learning based computer vision tasks such as single image super-resolution or image denoising, it is often more efficient to utilize patches of image training method instead of whole images. For instance, $\boldsymbol{x}^{(j)}$ and $\boldsymbol{y}^{(j)}$ can be patches of image from a ground truth data and a noisy image, respectively.

2.1.2 Stein's unbiased risk estimator (SURE)

Given the linear observation (II.1) and function $\boldsymbol{f}$, the SURE estimator for $\boldsymbol{f}(\boldsymbol{y})$ can be derived as:

$$\eta(\boldsymbol{f}(\boldsymbol{y}))=\frac{\|\boldsymbol{y}-\boldsymbol{f}(\boldsymbol{y})\|^2}{K}-\sigma^2+\frac{2\sigma^2}{K}\sum_{i=1}^{K}\frac{\partial\boldsymbol{f}_i(\boldsymbol{y})}{\partial\boldsymbol{y}_i} \tag{II.4}$$

where $\eta:\mathbb{R}^K\to\mathbb{R}$ and $\boldsymbol{y}_i$ is the ith element of $\boldsymbol{y}$. For a fixed $\boldsymbol{x}$, the following theorem holds:

Theorem 1 ([54, 55]). *The random variable $\eta(\boldsymbol{f}(\boldsymbol{y}))$ is an unbiased estimator of*

$$\mathrm{MSE}(\boldsymbol{f}(\boldsymbol{y}))=\frac{1}{K}\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y})\|^2$$

or

$$\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\left\{\frac{\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y})\|^2}{K}\right\}=\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\left\{\eta(\boldsymbol{f}(\boldsymbol{y}))\right\}$$

where $\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\{\cdot\}$ is the expectation operator dealing with the random vector $\boldsymbol{n}$, and MSE denotes the mean squared error. It should be note that in Theorem 1, $\boldsymbol{x}$ is coped wtih as a

deterministic, fixed vector. The variance σ^2 is assumed to be known and $\|\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{y})\|^2$ only needs the input $\boldsymbol{y}$ and the output of the estimator $\boldsymbol{f}(\boldsymbol{y})$. However, in the case of general DNN denoisers or estimators $\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})$, the divergence term presents the main computational challenge of calculating (II.4).

2.1.3 Monte-Carlo Stein's unbiased risk estimator (MC-SURE)

The Monte-Carlo (MC) method was introduced by Ramani *et al.* to estimate the divergence term in (II.4) for typical denoisers. For a fixed ground truth signal $\boldsymbol{x}$, the following theorem holds:

Theorem 2 ([56]). *Let $\tilde{\boldsymbol{n}} \sim \mathcal{N}(0,1) \in \mathbb{R}^K$ be independent of $\boldsymbol{n}$ and $\boldsymbol{y}$. Then,*

$$\sum_{i=1}^{K} \frac{\partial \boldsymbol{f}_i(\boldsymbol{y})}{\partial \boldsymbol{y}_i} = \lim_{\epsilon \to 0} \mathbb{E}_{\tilde{\boldsymbol{n}}} \left\{ \tilde{\boldsymbol{n}}^{\mathrm{t}} \left(\frac{\boldsymbol{f}(\boldsymbol{y} + \epsilon \tilde{\boldsymbol{n}}) - \boldsymbol{f}(\boldsymbol{y})}{\epsilon} \right) \right\} \tag{II.5}$$

provided that $\boldsymbol{f}(\boldsymbol{y})$ admits a well-defined second-order Taylor expansion. If not, this is still weakly valid when $\boldsymbol{f}(\boldsymbol{y})$ is tempered.

From the Theorem 2, the divergence term in (II.4) can be approximated by using a realization of $\tilde{\boldsymbol{n}} \sim \mathcal{N}(0,1)$ and a fixed small positive value ϵ:

$$\frac{1}{K} \sum_{i=1}^{K} \frac{\partial \boldsymbol{f}_i(\boldsymbol{y})}{\partial \boldsymbol{y}_i} \approx \frac{1}{\epsilon K} \tilde{\boldsymbol{n}}^{\mathrm{t}} \left(\boldsymbol{f}(\boldsymbol{y} + \epsilon \tilde{\boldsymbol{n}}) - \boldsymbol{f}(\boldsymbol{y}) \right) \tag{II.6}$$

where t is the transpose operator. Combining (II.6) with (II.4) allows the calculation of an unbiased estimator for the MSE, which can be practically used as a loss function to determine the network parameters $\boldsymbol{\theta}$ in the general denoisers $\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})$.

CHAPTER III

Methods

3.1 Methods

3.1.1 Unsupervised training of DNN-based Gaussian denoisers

Here, the work of [40] is revisited and a more rigorous derivation of the MC-SURE loss for the risk function in (II.2) is introduced. In addition, the uniform variance assumption in [40] is extended to the non-uniform variance for WSURE.

First, the original risk function in (II.2) is equivalent to the following function using conditioning:

$$\mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}\left[\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\left(\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2|\boldsymbol{x}\right)\right]. \tag{III.1}$$

Then, for a fixed $\boldsymbol{x}$,

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\left(\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2|\boldsymbol{x}\right) \\ &=\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2=K\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}_{0,\sigma^2}}\eta(\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta}))\end{aligned} \tag{III.2}$$

from Theorem 1. Thus, the original risk function is equivalent to the following modified risk function that does not require the separation of $\boldsymbol{x}$ and $\boldsymbol{n}$:

$$\mathbb{E}_{\boldsymbol{x},\boldsymbol{n}\sim p(\boldsymbol{x}),\mathcal{N}_{0,\sigma^2}}\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2=\mathbb{E}_{\boldsymbol{y}\sim q(\boldsymbol{y})}K\eta(\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})) \tag{III.3}$$

where $q(\boldsymbol{y})$ is the distribution function of $\boldsymbol{y}$ in (II.1) such that $\boldsymbol{x}, \boldsymbol{n} \sim p(\boldsymbol{x}), \mathcal{N}_{0,\sigma^2}$. It should be note that no clean ground truth data $\boldsymbol{x}$ are needed for the risk function on right-hand side of (III.3), and there is no approximation in (III.3). In other words, the original risk function in (II.2) with the ground truth is equal to the expected SURE term without the ground truth.

The SLLN allows (III.3) to be well-approximated as the following objective function, which is an unbiased estimator of (II.2):

$$\frac{1}{M}\sum_{j=1}^{M}\Bigg\{\|\boldsymbol{y}^{(j)}-\boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})\|^2 - K\sigma^2 + 2\sigma^2\sum_{i=1}^{K}\frac{\partial \boldsymbol{f}_i(\boldsymbol{y}^{(j)};\boldsymbol{\theta})}{\partial \boldsymbol{y}_i}\Bigg\}. \tag{III.4}$$

Finally, the last divergence term in (III.4) can be approximated using MC-SURE, with the unbiased function for (II.2) given by :

$$\frac{1}{M}\sum_{j=1}^{M}\Bigg\{\|\boldsymbol{y}^{(j)}-\boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})\|^2 - K\sigma^2 + \frac{2\sigma^2}{\epsilon}(\tilde{\boldsymbol{n}}^{(j)})^{\mathrm{t}}\left(\boldsymbol{f}(\boldsymbol{y}^{(j)}+\epsilon\tilde{\boldsymbol{n}}^{(j)};\boldsymbol{\theta}) - \boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})\right)\Bigg\}, \tag{III.5}$$

where ϵ is a small fixed positive number and $\tilde{\boldsymbol{n}}^{(j)}$ is a single variable from the standard normal distribution for each training data j. We randomly permuted the order of $\boldsymbol{y}^{(j)}$ and generated the new set of $\tilde{\boldsymbol{n}}^{(j)}$ in every epoch.

In [40], this final loss in (III.5) was optimized using a stochastic gradient-based optimization algorithm for training, such as the SGD or Adam optimization algorithms.

In this work, we modified our original signal model (II.1) to incorporate different noise levels at each pixel and for each sample as follows:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}_w, \qquad \boldsymbol{n}_w \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{III.6}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_K^2 \end{bmatrix}.$$

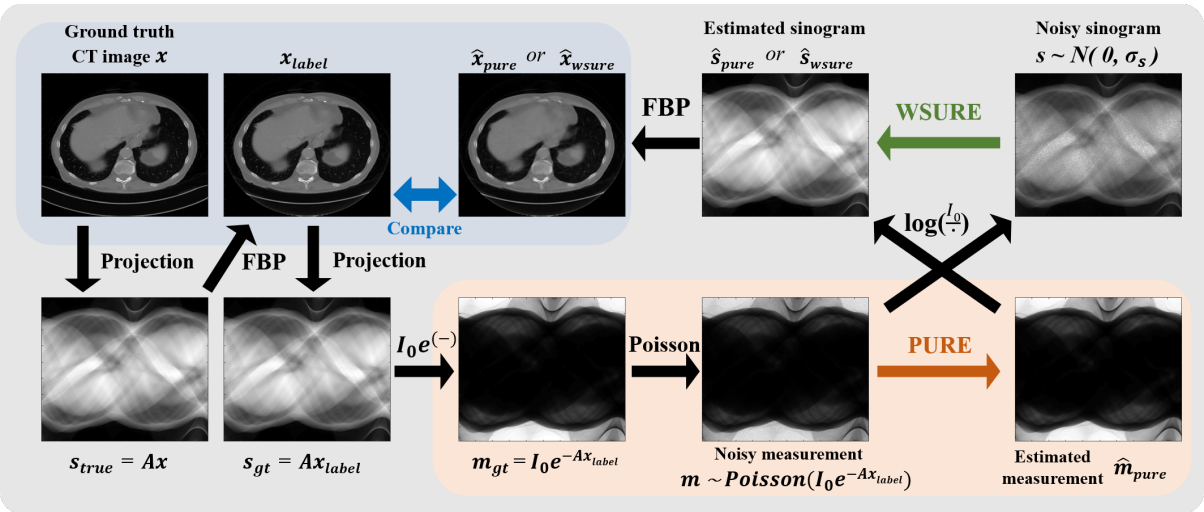


Figure 3.1: Overview of our proposed methods and details of the X-ray CT noise model. We proposed unsupervised training approaches for the DNN denoisers in two domains, namely the measurement and sinogram domains. As the data are corrupted by Poisson noise in the measurement domain, the PURE estimator can be used to predict the MSE without the ground truth measurement $\boldsymbol{m_{gt}}$. After computing logarithm of the measurement, the sinogram data can be approximated by weighted Gaussian noise and thus weighted SURE can be used to estimate the MSE without the ground truth sinogram $\boldsymbol{s_{gt}}$. Subsequently, the CT images can be reconstructed from the denoised measurement and sinogram to yield $\hat{\boldsymbol{x}}_{pure}$ and $\hat{\boldsymbol{x}}_{wsure}$, respectively. To evaluate our proposed methods, $\boldsymbol{x_{label}}$ was used as the target image.

Then, the final weighted MC-SURE (WSURE) loss can be derived as follows:

$$\frac{1}{M}\sum_{j=1}^{M}\Bigg\{\|\boldsymbol{y}^{(j)}-\boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})\|^2-\sum_{i=1}^{K}(\sigma_i^2)^{(j)} \\ +\frac{2}{\epsilon}(\boldsymbol{\Sigma}^{(j)}\tilde{\boldsymbol{n}}^{(j)})^{\mathrm{t}}\left(\boldsymbol{f}(\boldsymbol{y}^{(j)}+\epsilon\tilde{\boldsymbol{n}}^{(j)};\boldsymbol{\theta})-\boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})\right)\Bigg\}. \tag{III.7}$$

3.1.2 Unsupervised training of DNN based Poisson denoisers

In low-count imaging systems, the Poisson model is often used as a noise model:

$$\boldsymbol{y}=\zeta\boldsymbol{z} \;\text{ with }\; \boldsymbol{z}\sim Poisson(\boldsymbol{x}/\zeta) \tag{III.8}$$

where $\boldsymbol{z}\in\mathbb{R}^K$ is a random variable that follows the Poisson distribution and $\zeta\geq 0$ is the gain of the acquisition process that is related to the noise level.

Then, similar to (III.1), the risk function with Poisson noise can be decomposed into the

following two consecutive expectations using conditioning:

$$\mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}\left[\mathbb{E}_{\boldsymbol{y}\sim p(\boldsymbol{y}|\boldsymbol{x})}\left(\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2|\boldsymbol{x}\right)\right] \tag{III.9}$$

and the following inner term

$$\mathbb{E}_{\boldsymbol{y}\sim p(\boldsymbol{y}|\boldsymbol{x})}\left(\|\boldsymbol{x}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2|\boldsymbol{x}\right), \tag{III.10}$$

can be treated as the MSE for a fixed ground truth image $\boldsymbol{x}$ with $p(\boldsymbol{y}|\boldsymbol{x})$ as the Poisson noise model defined in (III.8).

Similar to the SURE estimator, the MC-PURE [47] can be used to approximate the MSE in (III.10):

$$\begin{aligned}&\|\boldsymbol{y}-\boldsymbol{f}(\boldsymbol{y};\boldsymbol{\theta})\|^2-\zeta\sum_{i=1}^{K}\boldsymbol{y}_i\\&+\frac{2\zeta}{\dot{\epsilon}}(\dot{\boldsymbol{n}}\odot\boldsymbol{y})^t\left(\boldsymbol{f}(\boldsymbol{y}+\dot{\epsilon}\dot{\boldsymbol{n}})-\boldsymbol{f}(\boldsymbol{y})\right)\end{aligned} \tag{III.11}$$

where $\dot{\boldsymbol{n}}\in\mathbb{R}^K$ is a random variable that follows a binary distribution taking values of -1 and 1 each with a probability of 0.5, $\dot{\epsilon}$ is a small positive number, and $\odot$ is an elementwise multiplication. It is assumed that the value of the noise parameter ζ is known. Thus, the final PURE loss for estimating (III.9) is given by:

$$\begin{aligned}&\frac{1}{M}\sum_{j=1}^{M}\Bigg\{\|\boldsymbol{y}^{(j)}-\boldsymbol{f}(\boldsymbol{y}^{(j)};\boldsymbol{\theta})\|^2-\zeta\sum_{i=1}^{K}\boldsymbol{y}_i^{(j)}\\&+\frac{2\zeta}{\dot{\epsilon}}(\dot{\boldsymbol{n}}^{(j)}\odot\boldsymbol{y}^{(j)})^t\left(\boldsymbol{f}(\boldsymbol{y}^{(j)}+\dot{\epsilon}\dot{\boldsymbol{n}}^{(j)})-\boldsymbol{f}(\boldsymbol{y}^{(j)})\right)\Bigg\}.\end{aligned} \tag{III.12}$$

3.1.3 Analytical X-ray CT reconstruction

A typical mono-energy X-ray CT measurement $\boldsymbol{m}$ generated from an attenuation image $\boldsymbol{x}$ can be modeled mathematically as:

$$\boldsymbol{m}\sim Poisson(I_0e^{-\boldsymbol{A}\boldsymbol{x}}+\boldsymbol{r}), \tag{III.13}$$

where $\boldsymbol{A}$ is a projection matrix that corresponds to the Radon transform for parallel-beam geometry (or other transforms for different geometries), and $\boldsymbol{x}$ is the ground truth CT image. I_0 denotes the X-ray source intensity that is usually determined by the tube current or voltage,

and $\boldsymbol{r}$ denotes the electrical noise and background contributions of scatter. Thus, low-dose CT corresponds to using a small I_0, while full-dose CT employs a large I_0. For a small I_0, $\boldsymbol{m}$ strictly follows the Poisson model, but for a large I_0, $\boldsymbol{m}$ is well-approximated with a Gaussian model. In the proposed Poisson noise model defined in (III.8), we set $\zeta = 1$. In this work, scatter in the measurement domain was not considered (*i.e.*, $\boldsymbol{r} = 0$).

Then, two steps are required to obtain the final reconstructed image $\hat{\boldsymbol{x}}$ from the measurement $\boldsymbol{m}$: 1) The measurement $\boldsymbol{m}$ is converted into the sinogram $\boldsymbol{s}$ or

$$\boldsymbol{s} = \log\left(I_0/\boldsymbol{m}\right), \tag{III.14}$$

and 2) the inverse Radon transform is applied using the filtered back projection (FBP) or

$$\hat{\boldsymbol{x}} = \boldsymbol{A}^{\dagger}\boldsymbol{s}$$

where $\boldsymbol{A}^{\dagger}$ is the pseudo-inverse matrix of $\boldsymbol{A}$. As the ramp filter of FBP in $\boldsymbol{A}^{\dagger}$ is a high-pass filter, denoising the measurement $\boldsymbol{m}$ or the sinogram $\boldsymbol{s}$ can greatly improve the quality of the reconstructed image $\hat{\boldsymbol{x}}$.

3.1.4 PURE based training of denoisers on CT measurements

For the ground truth image $\boldsymbol{x_{label}}$, DNNs can be trained to denoise the measurement $\boldsymbol{m}$ directly. In the supervised approach, these denoisers can be trained by minimizing the following MSE loss with the ground truth images:

$$\frac{1}{M}\sum_{j=1}^{M}\|\boldsymbol{f}(\boldsymbol{m}^{(j)};\boldsymbol{\theta}) - \boldsymbol{m_{gt}}^{(j)}\|^2 \tag{III.15}$$

where

$$\boldsymbol{m_{gt}} = I_0 e^{-\boldsymbol{A}\boldsymbol{x}_{label}}.$$

Here, we propose an unsupervised approach to train DNN denoisers without ground truth data using the proposed PURE loss defined in (III.12). Then, our proposed loss function is straightforward and is given by:

$$\begin{aligned}\frac{1}{M}\sum_{j=1}^{M}\Bigg\{&\|\boldsymbol{m}^{(j)} - \boldsymbol{f}(\boldsymbol{m}^{(j)};\boldsymbol{\theta})\|^2 - \sum_{i=1}^{K}\boldsymbol{m}_i^{(j)} \\ &+ \frac{2}{\dot{\epsilon}}(\dot{\boldsymbol{n}}^{(j)}\odot\boldsymbol{m}^{(j)})^t\left(\boldsymbol{f}(\boldsymbol{m}^{(j)}+\dot{\epsilon}\dot{\boldsymbol{n}}^{(j)}) - \boldsymbol{f}(\boldsymbol{m}^{(j)})\right)\Bigg\}.\end{aligned} \tag{III.16}$$

Note that (III.16) does not contain the ground truth. By minimizing this loss function with respect to $\boldsymbol{\theta}$, the DNN denoiser can be trained without the clean full-dose CT data. After the DNN denoiser is trained with this loss function defined in (III.16), the estimated denoised measurement data $\hat{\boldsymbol{m}}_{\boldsymbol{pure}}$ can be obtained using

$$\hat{\boldsymbol{m}}_{\boldsymbol{pure}} = \boldsymbol{f}(\boldsymbol{m}; \boldsymbol{\theta}).$$

To obtain a denoised CT image from the estimated measurement data $\hat{\boldsymbol{m}}_{\boldsymbol{pure}}$, the measurement $\hat{\boldsymbol{m}}_{\boldsymbol{pure}}$ is first converted into the sinogram $\hat{\boldsymbol{s}}_{\boldsymbol{pure}}$ and then FBP is applied to $\hat{\boldsymbol{s}}_{\boldsymbol{pure}}$ to obtain the final reconstructed image using

$$\hat{\boldsymbol{x}}_{\boldsymbol{pure}} = \boldsymbol{A}^{\dagger}\hat{\boldsymbol{s}}_{\boldsymbol{pure}}.$$

Fig. 3.1 illustrates the pipeline of the proposed methods.

3.1.5 WSURE based training of denoisers on CT sinograms

We also investigated DNN denoisers on CT sinograms. The noise of the sinogram obtained in (III.14) can be approximately modeled as a weighted Gaussian distribution [48]:

$$\boldsymbol{s}_i = \boldsymbol{s}_{\boldsymbol{gt}} + \boldsymbol{n}_{\boldsymbol{s}}, \quad \boldsymbol{n}_{\boldsymbol{s}} \sim \mathcal{N}(0, \sigma_s^2) \ \text{ where } \sigma_s = \sqrt{\boldsymbol{m}_i}/I_0 \tag{III.17}$$

where $\boldsymbol{s}_{\boldsymbol{gt}} = \boldsymbol{A}\boldsymbol{x}_{\boldsymbol{label}}$, $\boldsymbol{s}_i$ is the i^{th} element of the noisy sinogram vector, and σ_s is the standard deviation of the weighted Gaussian model in the sinogram domain. Note that the accuracy of the sinogram noise model in (III.17) depends on the value of I_0. For example, for a very small I_0, the model in (III.17) did not yield good approximation results in our simulations. However, for a modestly large I_0, the model in (III.17) was fairly accurate due to the central limit theorem.

Training a DNN denoiser in a supervised way is equivalent to minimizing the following MSE:

$$\frac{1}{M}\sum_{j=1}^{M} \|\boldsymbol{f}(\boldsymbol{s}^{(j)}; \boldsymbol{\theta}) - \boldsymbol{s_{gt}}^{(j)}\|^2 \tag{III.18}$$

Herein, we propose an unsupervised training method that utilizes the following SURE loss function without ground truth images to unbiasedly estimate (III.18):

$$\frac{1}{M}\sum_{j=1}^{M} \left\{ \|\boldsymbol{s}^{(j)} - \boldsymbol{f}(\boldsymbol{s}^{(j)}; \boldsymbol{\theta})\|^2 - \sum_{i=1}^{K} (\sigma_s^2)_i^{(j)} \right. \tag{III.19}$$

$$+\frac{2}{\epsilon}(\boldsymbol{\sigma_s^2}^{(j)}\odot\tilde{\boldsymbol{n}}^{(j)})^{\mathrm{t}}\left(\boldsymbol{f}(\boldsymbol{s}^{(j)}+\epsilon\tilde{\boldsymbol{n}}^{(j)};\boldsymbol{\theta})-\boldsymbol{f}(\boldsymbol{s}^{(j)};\boldsymbol{\theta})\right)\Big\}$$

where $\boldsymbol{\sigma_s^2}$ is a vector whose i^{th} element is equal to $(\sigma_s^2)_i$.

Often, iterative algorithms for low-dose CT reconstructions minimize the weighted MSE instead of the MSE defined in (III.18) [7, 8]. We investigated the weighted MSE (WMSE) as the loss function with ground truth images and its corresponding unsupervised loss function. The WMSE loss that considers noise statistics [48] is given by:

$$\frac{1}{M}\sum_{j=1}^{M}\|\boldsymbol{f}(\boldsymbol{s}^{(j)};\boldsymbol{\theta})-\boldsymbol{s_{gt}}^{(j)}\|_{\boldsymbol{W}}^2 \tag{III.20}$$

where $\boldsymbol{W}$ is a diagonal matrix whose i^{th} diagonal element is $\boldsymbol{m}_i$ or $\boldsymbol{W}_{ii}=\boldsymbol{m}_i$. From $\boldsymbol{s_i}$, a weight matrix can be obtained as follows:

$$\boldsymbol{m}_i = I_0\boldsymbol{e}^{-\boldsymbol{s}_i}.$$

For this supervised loss function with the ground truth, we propose the following weighted SURE (WSURE) that can provide an unbiased estimate of (III.20):

$$\begin{aligned}\frac{1}{M}\sum_{j=1}^{M}\Big\{&\|\boldsymbol{s}^{(j)}-\boldsymbol{f}(\boldsymbol{s}^{(j)};\boldsymbol{\theta})\|_{\boldsymbol{W}}^2-\sum_{i=1}^{K}\boldsymbol{W}_{ii}(\sigma_s^2)_i^{(j)}\\&+\frac{2}{\epsilon}(\boldsymbol{W}\boldsymbol{\sigma_s^2}^{(j)}\odot\tilde{\boldsymbol{n}}^{(j)})^{\mathrm{t}}\left(\boldsymbol{f}(\boldsymbol{s}^{(j)}+\epsilon\tilde{\boldsymbol{n}}^{(j)};\boldsymbol{\theta})-\boldsymbol{f}(\boldsymbol{s}^{(j)};\boldsymbol{\theta})\right)\Big\}.\end{aligned} \tag{III.21}$$

To reconstruct the CT images from the denoised sinograms, first, the estimated sinograms are obtained from the trained DNN denoisers with the loss function defined in (III.21) as follows:

$$\hat{\boldsymbol{s}}_{\boldsymbol{wsure}}=\boldsymbol{f}(\boldsymbol{s};\boldsymbol{\theta})$$

By applying the FBP method, the reconstructed CT image $\hat{\boldsymbol{x}}_{\boldsymbol{wsure}}$ is obtained:

$$\hat{\boldsymbol{x}}_{\boldsymbol{wsure}}=\boldsymbol{A}^{\dagger}\hat{\boldsymbol{s}}_{\boldsymbol{wsure}}.$$

Fig. 3.1 illustrates the pipeline of the proposed methods.

CHAPTER IV

Experiments and Results

4.1 Experiments and Results

In this section, the denoising experiment results are showed with the AAPM CT image dataset from the 2016 Low-Dose CT Grand Challenge [34] using a deep convolutional neural network (CNN) image denoiser, which is a altered version of U-Net [2]. The modified U-Net has been still used to achieve state-of-the-art denoising performance [50] and appears to be a good candidate for presenting the performance of unsupervised training methods [39]. All of the neural networks described in this work (denoted by NET, which can refer to U-Net) were trained with one of the following optimization objectives: (MSE) the minimum MSE between an estimated signal and its ground truth signal in (II.3) and (SURE, WSURE, PURE) the minimum corresponding empirical risk without the ground truth in (III.12), (III.19), and (III.21).

Table 4.1: Summary of various denoising methods. NET can refer to SDA / U-NET.

Method	Description
BM3D [43]	Conventional state-of-the-art method
DIP [35]	State-of-the-art unsupervised method
Deep Decoder [38]	State-of-the-art unsupervised method
NET-SURE	Optimizing SURE without ground truth
NET-WSURE	Optimizing WSURE without ground truth
NET-PURE	Optimizing PURE without ground truth
NET-EST-FT	Optimizing EST (estimator) without ground truth, by fine-tuning on noisy test data
NET-MSE-GT	Optimizing MSE with ground truth

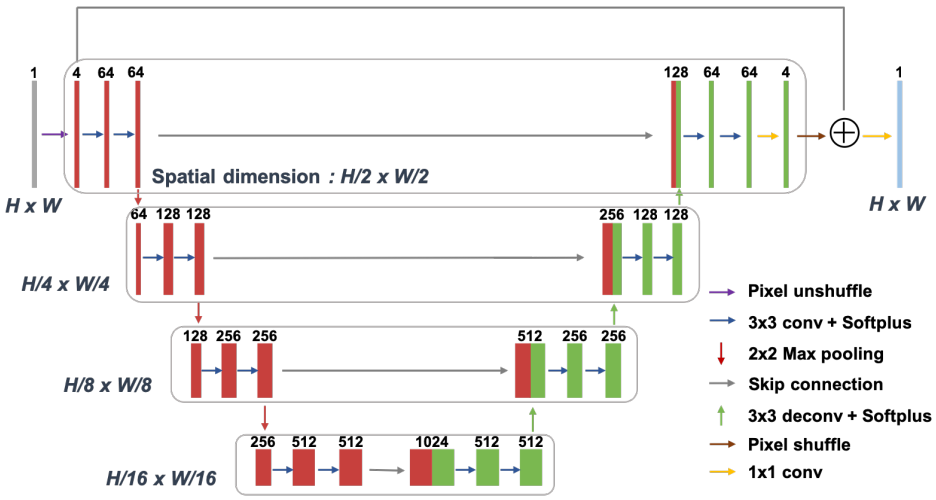


Figure 4.1: The network architecture of the proposed method, which is a modified version of U-Net [2]. The pixel unshuffle and shuffle layers reshape the data to reduce the memory usage of the GPU. Differentiable activation functions such as Softplus were used without loss of performance as suggested in [3].

The NET-MSE methods generated noisy training images at every epoch as similar to [49], while the proposed NET-SURE, NET-PURE, and NET-WSURE, only used the noisy images obtained before training. These methods are denoted by EST, which can refer to SURE, PURE or WSURE. We also proposed the EST-FT method which only utilizes a single test data for fine-tuning (refining) with the pre-trained network. Table 4.1 shows all the simulated configurations, including conventional state-of-the-art denoisers, the BM3D [43], DIP [35], and Deep Decoder [38] that did not need any training and ground truth data.

4.1.1 Low-dose CT reconstruction with the AAPM CT dataset

Poisson denoising experiments were performed with the AAPM CT dataset [34]. This dataset contains 2,378 512 × 512 CT images with 3mm slice thickness, and full / quarter doses from 10 patients. To generate the training and testing image pairs, the full-dose CT images were set as ground truth and Poisson noise was added to the ground truth measurement. The CT images of 8 patients were assigned to the training set, 1 to the validation set, and 1 to the test set. For training, the horizontal flip and flop data augmentations were adopted.

The network architecture used in this experiment is based on the modified version of U-Net [2], as shown in Fig 4.1. It consists of a contracting part and an expansive part. The contracting part consists of the repetitive application of two 3×3 convolutions, each followed by Softplus [3] and a 2×2 max pooling operation with stride 2 for downsampling. To approximate

Table 4.2: Comparison results for supervised training of DNN denoisers with MSE in various domains. Averages of the PSNR (Performance in dB) and SSIM values are reported.

Domain	Method	$I_0 = 10^3$	$I_0 = 10^4$	$I_0 = 10^5$
Measurement	U-net-MSE-GT	41.787/0.9826	46.725/0.9889	49.684/0.9960
Sinogram	U-net-MSE-GT	43.373/0.9866	47.583/0.9928	51.340/0.9963
Sinogram	U-net-WMSE-GT	42.552/0.9845	46.050/0.9905	49.350/0.9939
Image	U-net-MSE-GT	39.942/0.9692	44.258/0.9825	48.686/0.9924

the MC divergence term in the proposed unsupervised loss, differentiable activation functions are more effective than non-differentiable activation functions such as the ReLU and PReLU. This was first observed in [3] for SURE, and the same phenomena for PURE and WSURE were also observed in this work. The expansive path is similar to the contracting part. It differs from the contracting part in that 3×3 deconvolutions are used for upsampling.

The measurement and sinogram data should be used as global information, rather than as local information. In this case, the patch extraction method for training does not seem appropriate for denoising in the measurement and sinogram domains. Following data transformation from the image domain to the measurement or sinogram domains, the size of the measurement or sinogram data was 888 × 984. The large data size results in an inefficient GPU memory access. To address this memory issue, unshuffle and shuffle layer were employed into top and bottom of the network for each pixel. In all the cases, the network was trained with the Adam optimizer with a learning rate of 10^{-4} for 100 epochs. Then, the learning rate was decreased by half for the next 100 epochs. The batch size was set to 4. The ϵ value was set to 10^{-4} in all the cases. All experiments were implemented with PyTorch for the denoising simulations and with MATLAB and Michigan Image Reconstruction Toolbox[1] for the analytical CT reconstructions. Using a single NVIDIA Titan Xp GPU, training took ∼10 hours for U-net-MSE-GT and approximately ∼20 hours for U-net-PURE and U-net-WSURE.

4.1.2 Oracles: Supervised low-dose CT reconstructions

First, we compared different denoising methods in the image, sinogram, and measurement domains for low-dose CT reconstructions when training DNN denoisers in supervised ways with ground truth data. These results will be used as oracles for our unsupervised trainings in the low-dose CT reconstructions in sinogram / measurement domains. All results are summarized in Table 4.2.

It was observed that the simulated results in the image domain yielded relatively poor performance in terms of the PSNR and structural similarity (SSIM) in most of cases. Note that the

[1] Jeffrey A Fessler, available at https://web.eecs.umich.edu/ fessler/code/

denoisers in the image domain for CT reconstruction must address spatially dependent, colored noises unlike other methods in the sinogram / measurement domains that have independent noises. Thus, as addressing spatially dependent noise appears challenging compared with independent noise, the best performance was not able to be achieved. All denoising methods in the measurement / sinogram domains yielded comparable performances and U-net-MSE-GT with sinograms yielded better performance than the other methods with different losses and/or domains.

Note that our unsupervised methods aimed to train DNN denoisers that achieve as close a performance as possible to these oracle cases. High-performance oracles were one of the factors that determined the performance of our proposed unsupervised training methods.

4.1.3 Unsupervised low-dose CT reconstruction from denoised measurements

To obtain the estimated images, the outputs of DNN denoisers with the measurements were first converted into denoised sinograms, and then were reconstructed using FBP. The quantitative results obtained with the denoised measurements are described in Table 4.3. To evaluate the PSNR and SSIM of the output images, the full-dose image $\boldsymbol{x}_{\boldsymbol{label}}$ was set as the target image. While BM3D+VST [57] yielded good quantitative results, they were lower than the oracle cases with MSE based supervised training methods. Our proposed PURE based unsupervised training methods outperformed BM3D+VST in all the cases and our proposed PURE-FT based methods yielded consistently better performance than the PURE based methods in all the cases, with comparable performance to the MSE based methods.

The visual low-dose CT reconstruction results with the denoised measurements are shown in Fig. 4.2. All methods yielded substantially denoised images in all the cases. Our PURE-FT based method yielded the best root mean square error (RMSE) among all the compared methods in all cases. The DNN based methods seemed to perform well as compared with BM3D+VST [57] in the challenging extremely low-dose case, $I_0 = 10^3$, in terms of the RMSE performance and

Table 4.3: Quantitative results of the low-dose CT reconstructions from denoised measurements using various methods. Averages of the PSNR (Performance in dB) and SSIM values are reported.

Methods	$I_0 = 10^3$	$I_0 = 10^4$	$I_0 = 10^5$
BM3D+VST [57]	41.127/0.9657	46.328/0.9875	49.125/0.9927
U-net-PURE (22)	41.677/0.9772	46.631/0.9893	49.315/0.9937
U-net-PURE-FT	**41.712/0.9780**	**46.652/0.9895**	**49.412/0.9939**
U-net-MSE-GT (III.15)	41.787/0.9826	46.725/0.9889	49.684/0.9960

visual quality of the residual images (ideally, residuals images should have pixel values equal to zero).

4.1.4 Unsupervised low-dose CT reconstruction from denoised sinograms

For the sinogram domain experiments, the visual qualitative results of the denoised sinogram obtained with various methods are shown in Fig. 4.3. In the residual images of the BM3D and Deep Decoder, substantial noise was still observed after denoising, while strong structural patterns were observed in the residual images of the DIP. However, the DNN-based methods with both supervised training and our proposed unsupervised training achieved significantly reduced noise and structural patterns as compared with the other methods visually. These visual observations were well-reflected in the RMSE values. Note that our proposed WSURE and WSURE-FT methods had the lowest RMSE values among all methods without the ground truth, despite using an approximate noise model for the low-dose data with $I_0 = 10^4$.

To generate the estimated CT image, the sinogram outputs of the network were reconstructed by using FBP. Table 4.4 shows the quantitative results of our unsupervised training of sinogram denoisers for the low-dose CT reconstruction. For all counts, our unsupervised SURE / WSURE training of DNN sinogram denoisers did not yield comparable quantitative results to the supervised DNN sinogram denoisers as compared with our PURE based methods. This can possibly attributed to the inaccurate approximation of the noise model in the sinogram domain. These performance gaps between the supervised and unsupervised methods were especially large for low counts, such as for $I_0 = 10^3$ and $I_0 = 10^4$. However, SURE / WSURE also presented many advantages. First, the noise approximation is more accurate for higher counts such as $I_0 = 10^5$ due to the central limit theorem. Second, fine-tuning of a single test image works better than in the case of PURE-based methods and yields substantially improved performance, as shown in Table 4.4. Due to these advantages, our proposed U-net-WSURE-FT yielded the best performance for $I_0 = 10^5$ among all methods without using the ground truth.

Table 4.4: Quantitative results of the low-dose CT reconstructions from denoised sinograms using various methods. Averages of the PSNR (Performance in dB) and SSIM values are reported.

Methods	$I_0 = 10^3$	$I_0 = 10^4$	$I_0 = 10^5$
BM3D [43]	16.772/0.2026	29.929/0.6786	48.668/0.9932
U-net-SURE (III.19)	32.886/0.7818	40.365/0.9535	47.120/0.9891
U-net-WSURE (III.21)	37.260/0.9269	42.811/0.9751	48.828/0.9934
U-net-WSURE-FT	**37.352/0.9273**	**42.948/0.9753**	**49.423/0.9942**
U-net-MSE-GT (III.18)	43.373/0.9866	47.583/0.9928	51.340/0.9959
U-net-WMSE-GT (III.20)	42.552/0.9845	46.050/0.9905	49.350/0.9939

Table 4.5: Visual results of the low-dose CT reconstructions from denoised sinograms using various methods on 10 randomly selected images (Performance in dB).

Image	1	2	3	4	5	6	7	8	9	10
					$I_0 = 10^3$					
BM3D [43]	21.39	21.45	21.44	21.07	21.01	18.66	17.92	16.58	16.03	15.46
DIP [35]	35.01	24.45	34.64	35.34	26.56	35.64	31.99	34.17	29.05	22.66
DeepDecoder [38]	30.81	31.53	30.60	30.32	30.01	28.53	28.52	28.25	28.66	28.15
U-net-WSURE	38.49	38.45	38.62	38.62	38.34	37.93	37.71	37.16	36.92	36.60
U-net-WSURE-FT	**38.67**	**38.61**	**38.82**	**38.87**	**38.51**	**38.13**	**37.79**	**37.20**	**36.94**	**36.60**
U-net-WMSE-GT	42.34	42.32	42.73	42.83	42.58	42.32	42.39	42.28	42.37	42.20
					$I_0 = 10^4$					
BM3D [43]	36.28	36.63	36.41	35.99	34.96	32.40	31.28	29.84	29.82	28.36
DIP [35]	38.32	31.22	34.61	37.58	34.61	35.71	37.45	39.71	38.52	36.52
DeepDecoder [38]	31.62	31.56	31.74	32.32	31.73	31.40	31.11	28.63	30.85	29.96
U-net-WSURE	43.84	43.81	43.99	44.03	43.78	43.42	43.29	42.74	42.59	42.31
U-net-WSURE-FT	**43.99**	**44.00**	**44.15**	**44.21**	**43.91**	**43.54**	**43.39**	**42.84**	**42.70**	**42.42**
U-net-WMSE-GT	45.34	45.27	45.66	45.85	45.60	45.41	45.43	45.30	45.46	45.36
					$I_0 = 10^5$					
BM3D [43]	48.61	48.57	48.81	49.02	48.93	49.06	49.21	48.88	48.91	48.85
DIP [35]	39.21	33.82	41.22	40.10	37.40	35.75	37.93	40.67	39.25	38.79
DeepDecoder [38]	31.82	32.50	32.12	32.14	29.94	29.75	29.65	30.52	31.17	30.83
U-net-WSURE	48.22	48.19	48.49	48.59	48.36	48.10	48.07	47.79	47.7	47.61
U-net-WSURE-FT	**49.83**	**49.84**	**50.10**	**50.16**	**49.82**	**49.36**	**49.23**	**48.73**	**48.64**	**48.40**
U-net-WMSE-GT	49.04	48.99	49.35	49.46	49.16	48.83	48.82	48.54	48.62	48.49

Table 4.5 also shows the comparison results obtained with the DIP and Deep Decoder for 10 randomly selected images (due to the slow reconstruction speed in the case of the DIP, ∼40 minutes were required per image). To simulate the DIP and Deep Decoder, the max iterations were set to 10,000 and 2,000, respectively. These results are also consistent with our proposed unsupervised methods in the sinogram domain yielding the best performance among all other methods such as the BM3D, DIP and Deep Decoder.

Fig. 4.4 illustrates the visual results of the denoising experiments in the sinogram domain with other state-of-the-art methods. For $I_0 = 10^5$, all methods except for the DIP and Deep-Decoder yielded comparably sharp images. However, for $I_0 = 10^4$, our proposed unsupervised WSURE methods yielded results comparable to the supervised MSE methods, and better than those of the BM3D, DIP and Deep Dedcoder. For $I_0 = 10^3$ our proposed methods yielded mildly noisy images, but still achieved the best image quality among other methods without the ground truth.

Fig. 4.5 illustrates the visual reconstruction results of WSURE-FT and PURE-FT in the coronal and sagittal directions for $I_0 = 10^4$. Both the WSURE-FT and PURE-FT methods

produced significantly denoised images. The WSURE-FT method generated a sharper image while PURE-FT method had a smaller RMSE value compared with each other.

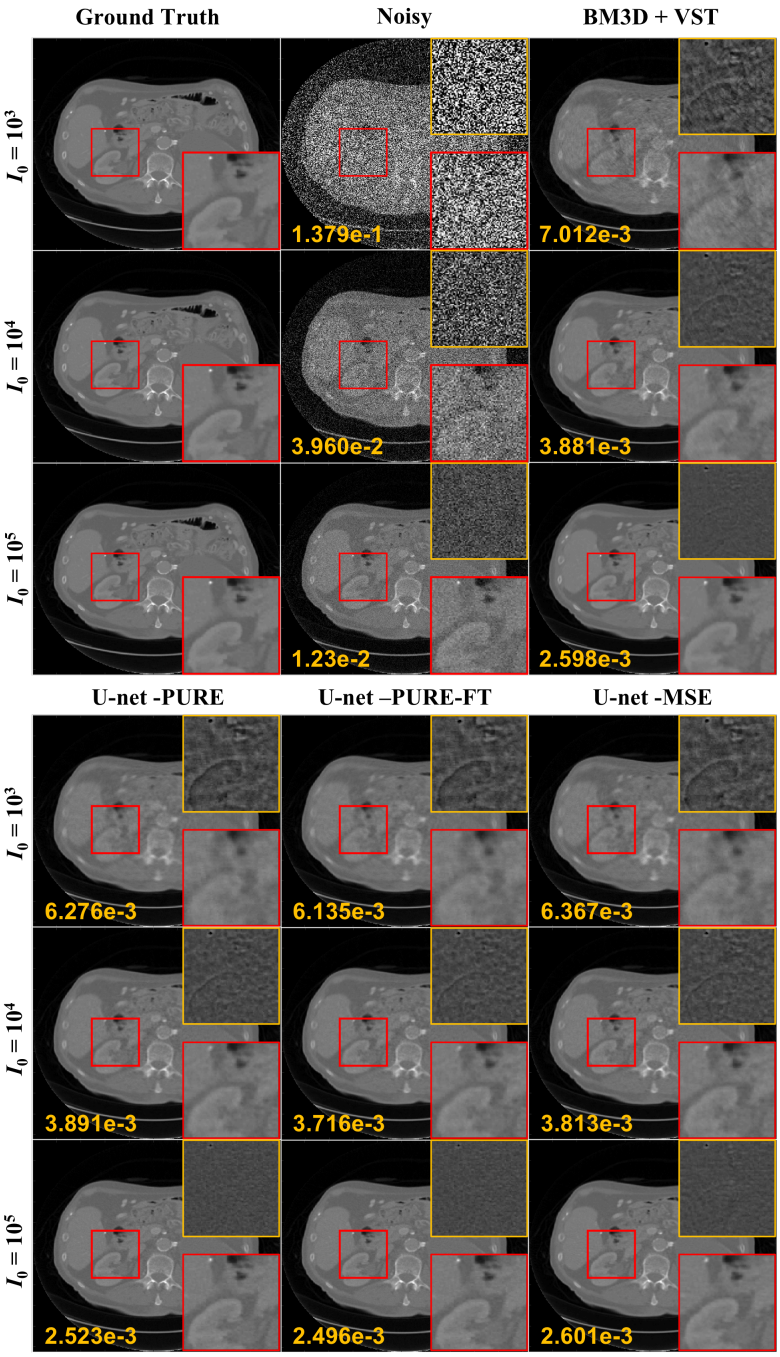


Figure 4.2: Visual results of the low-dose CT reconstructions from denoised measurements obtained with various methods. The results of DIP and Deep decoder were not included due to unstable training and poor performance. The red and yellow boxes represent the enlarged views and their corresponding residuals, respectively. The numbers in orange in the images indicate the root mean square error (RMSE) values. The intensity ranges were (-1000, 2000)[HU] for the CT images and (-200,200)[HU] for the residual images.

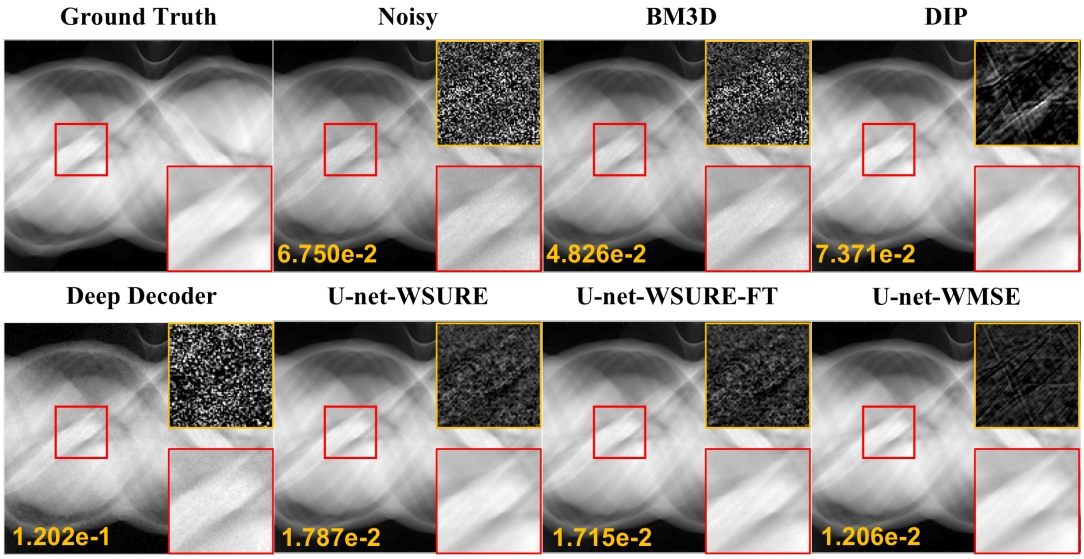


Figure 4.3: Visual results of the denoised sinograms obtained using various methods for $I_0 = 10^4$. The red and yellow boxes represent the enlarged views and their corresponding residuals, respectively. The numbers in orange in the images indicate the RMSE values. The intensity ranges were (-1000, 3000) and (-100,150) for the sinograms and residuals, respectively.

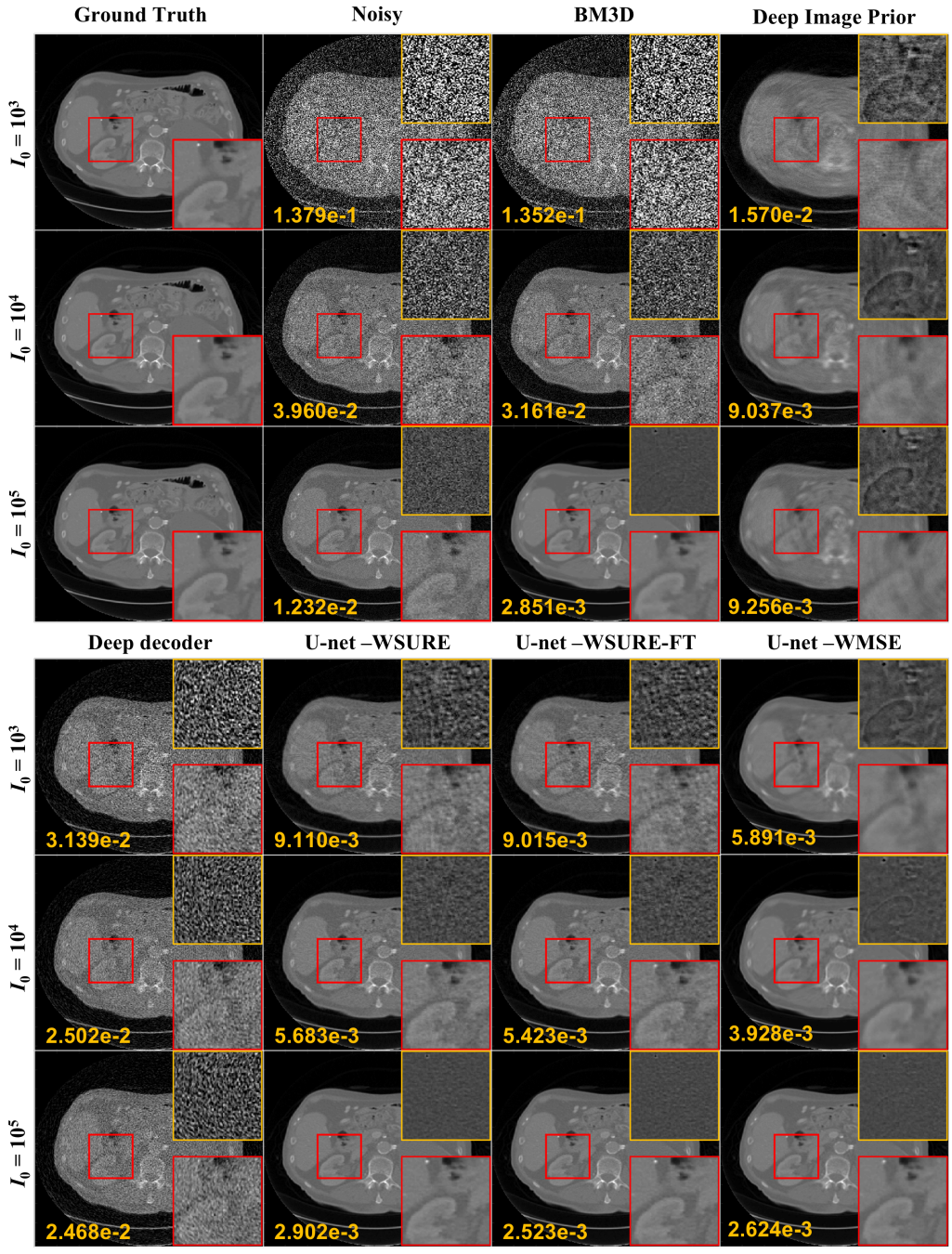


Figure 4.4: Visual results of the low-dose CT reconstructions from denoised sinograms obtained with various methods. The red and yellow boxes represent the enlarged views and their corresponding residuals, respectively. The numbers in orange in the images indicate the RMSE values. The intensity ranges were (-1000, 2000)[HU] and (-200,200)[HU] for the CT and residual images, respectively.

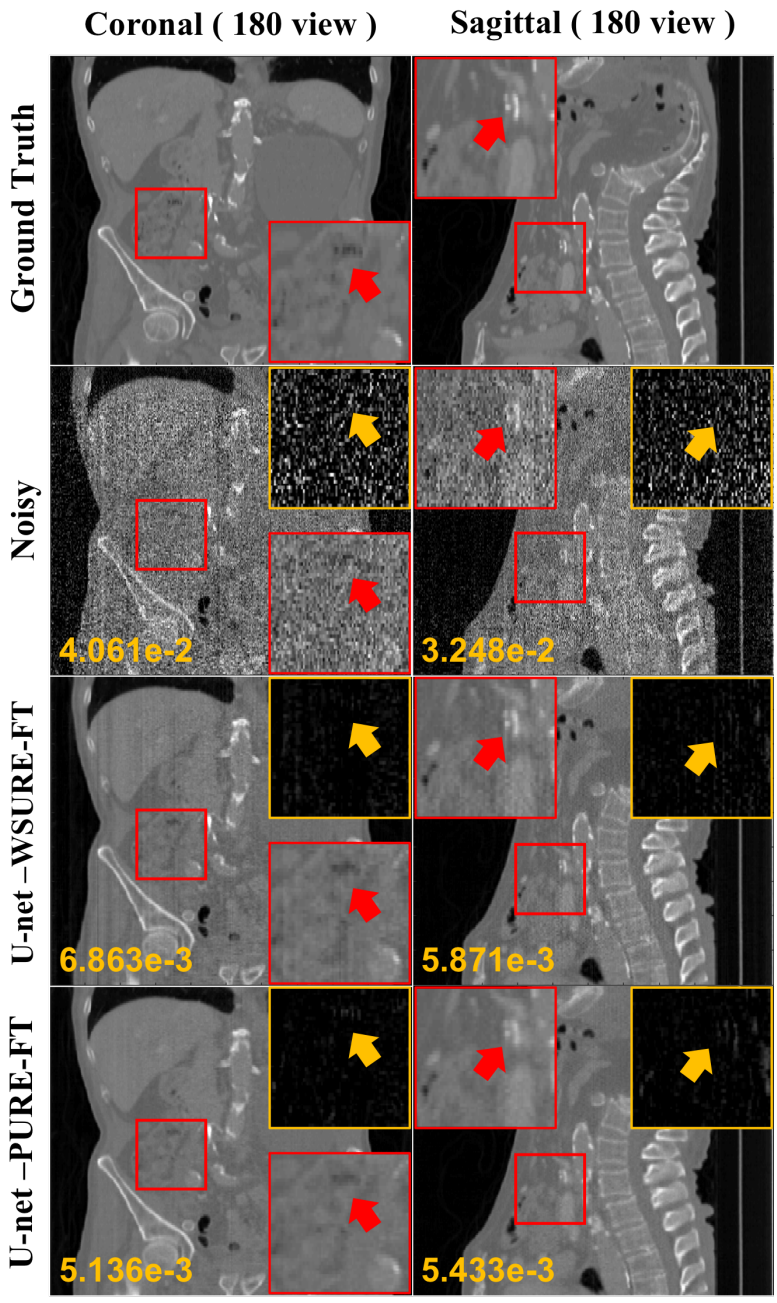


Figure 4.5: Coronal and sagittal views of the low-dose CT reconstruction from denoised sinograms obtained with various methods for $I_0 = 10^4$. The red and yellow boxes represent the enlarged views and their corresponding residuals, respectively. The numbers in orange in the images indicate the RMSE values. The intensity ranges were (-1000, 2000)[HU] and (-100,100)[HU] for the CT and residual images, respectively.

CHAPTER V

Analysis on proposed method

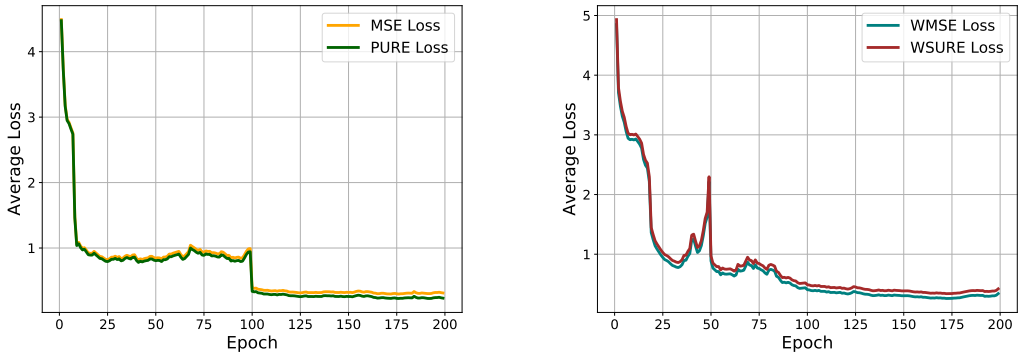
5.1 Analysis on proposed method

5.1.1 Training convergence of PURE and WSURE

Fig. 5.1 shows the training loss curves of our proposed PURE and WSURE for $I_0 = 10^5$. Due to the exact noise modeling of Poisson noise, the PURE loss function generated an accurate approximation loss curve compared with the MSE loss. From 100 epochs, the learning rate was reduced by half, which reduced the average loss. WSURE-based optimization also produced accurate approximation compared with WMSE loss over all epochs. It was empirically determined that objective function of both PURE and WSURE can be converged as iteration continues. These similar tendencies of PURE and WSURE compared to their oracle objective functions enabled the DNNs to be trained with the direction of minimizing the MSE or WMSE.

5.1.2 Computation time

Table 5.1 shows a comparison of the computation time required to denoise one measurement or sinogram data. The DIP and Deep decoder require 2199.77 seconds and 115.53 seconds per data, respectively. BM3D and BM3D+VST were much faster than the DIP and Deep decoder and took 1.56 seconds and 2.35 seconds, respectively, per sample. Our proposed method including the PURE and WSURE methods only required 0.06 second per sample, while the



(a) PURE (b) WSURE

Figure 5.1: Loss curves for the training of proposed network with (a) PURE and its corresponding MSE and (b) WSURE and its corresponding WMSE.

PURE method achieved the best performance in all the cases and the WSURE method achieved comparable performance to the other methods for $I_0 = 10^5$. Although our proposed methods with fine-tuning took 7.08 seconds per sample, improved performance was achieved.

Table 5.1: Computation time required to denoise one sample data (888×984).

Method	Time (sec)
BM3D [43]	1.56
BM3D+VST [57]	2.35
DIP [35]	2199.77
DeepDecoder [38]	115.53
Ours	**0.06**
Ours with fine tuning	7.08

CHAPTER VI

Discussion

6.1 Discussion

While many research studies have investigated DNN-based low-dose CT reconstructions with full-dose ground truth data / images, unsupervised training methods for DNNs in low-dose CT have been rarely studied. In this work, we explored two potential unsupervised training methods for low-dose CT in the measurement and sinogram domains.

PURE in the CT measurement domain appears to yield excellent estimation results for the MSE in the case of Poisson noise. Due to accurate noise modeling, our proposed unsupervised PURE methods produced comparable training results to the supervised MSE method. However, there are some limitations: 1) In the experimental setting, the noise level was controlled by the count I_0 defined in (III.13), and not by ζ defined in (III.8). To calculate the PURE loss in (III.11), ζ was set to 1. For higher values of $\zeta > 0.2$, the accuracy of the MSE loss approximation decreased. 2) Fine-tuning with PURE did not work well as compared with the case of WSURE methods. Despite these limitations, the PURE methods yielded state-of-the-art performance for all counts, as evidenced by the quantitative results, for the unsupervised training algorithms.

WSURE in the CT sinogram domain appears to yield excellent low-dose CT reconstructions because 1) at higher I_0, as the noise property had an accurate Gaussian distribution, the more stable estimator SURE was used to address σ compared to the PURE estimator for dealing with ζ and 2) fine-tuning worked with substantial performance improvement. However, there

remains rooms for further improvements in the WSURE methods. The most critical performance improvement can be achieved by developing more accurate noise modeling and its corresponding unbiased MSE estimation. Trade-offs exist between our proposed PURE and WSURE methods in the measurement and sinogram domains, respectively. In addition, the goal of this study was to estimate the results of oracle so that we can obtain the better results when enhanced network architecture is utilized for both methods. While denoising in the image domain with unsupervised DNN training was not studied in this work, it is certainly an interesting topic to pursue. The GSURE [58] or Bayesian approach [59] have a great potential as an unsupervised training methods in the image domain. However, note that its oracle performance in the image domain is not as good as that in the sinogram domain. Better deep neural networks that address complicated noise structures should be investigated in future studies.

CHAPTER VII

Conclusion

7.1 Conclusion

In this study, we investigated low-dose CT reconstruction methods with denoised sinograms and measurements using DNNs. While most approaches use supervised training with ground truth images, in this work, unsupervised training methods without full-dose CT images were proposed. Our PURE-based methods with accurate noise model in the measurement domain and our WSURE-based methods with approximate noise model in the sinogram domain outperformed all other state-of-the-art methods without the ground truth, such as the BM3D, DIP and DeepDecoder. The proposed methods also achieved comparable performance to the supervised methods with ground truth data in some cases.

References

[1] J.A. Fessler, *Fundamentals of CT Reconstruction in 2D and 3D*, pp. 263–295, 12 2014. iv, 3

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015, pp. 234–241. iv, 17, 18

[3] S. Soltanayev, R. Giryes, S. Y. Chun, and Y. C. Eldar, "On divergence approximations for unsupervised training of deep denoisers based on Stein's unbiased risk estimator," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 3592–3596. iv, 18, 19

[4] Emil Y Sidky and Xiaochuan Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine & Biology*, vol. 53, no. 17, pp. 4777–4807, 2008. 4

[5] Sathish Ramani and Jeffrey A Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE transactions on medical imaging*, vol. 31, no. 3, pp. 677–688, 2011. 4

[6] Marcel Beister, Daniel Kolditz, and Willi A Kalender, "Iterative reconstruction methods in X-ray CT," *Physica Medica: European Journal of Medical Physics*, vol. 28, no. 2, pp. 94–108, 2012. 4

[7] H. Nien and J. A. Fessler, "Fast x-ray ct image reconstruction using a linearized augmented lagrangian method with ordered subsets," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 388–399, 2015. 4, 16

[8] Donghwan Kim, Sathish Ramani, and Jeffrey A Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction," *IEEE transactions on medical imaging*, vol. 34, no. 1, pp. 167–178, 2015. 4, 16

[9] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser, "Deep Convolutional Neural Network for Inverse Problems in Imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, July 2017. 5

[10] H. Chen, Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou, and G. Wang, "LEARN: Learned Experts' Assessment-Based Reconstruction Network for Sparse-Data CT," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1333–1347, June 2018. 5

[11] Shipeng Xie, Xinyu Zheng, Yang Chen, Lizhe Xie, Jin Liu, Yudong Zhang, Jingjie Yan, Hu Zhu, and Yining Hu, "Artifact Removal using Improved GoogLeNet for Sparse-view CT Reconstruction," *Scientific Reports*, vol. 8, no. 6700, Apr 2018. 5

[12] Y. Han and J. C. Ye, "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1418–1429, June 2018. 5

[13] Hu Chen, Yi Zhang, Mannudeep K Kalra, Feng Lin, Yang Chen, Peixi Liao, Jiliu Zhou, and Ge Wang, "Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, Nov. 2017. 5

[14] Eunhee Kang, Junhong Min, and Jong Chul Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Medical Physics*, vol. 44, no. 10, pp. e360–e375, Oct. 2017. 5

[15] Jelmer M Wolterink, Tim Leiner, Max A Viergever, and Ivana Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2536–2545, 2017. 5

[16] Qingsong Yang, Pingkun Yan, Yanbo Zhang, Hengyong Yu, Yongyi Shi, Xuanqin Mou, Mannudeep K Kalra, Yi Zhang, Ling Sun, and Ge Wang, "Low-dose CT image denoising

using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1348–1357, 2018. 5

[17] E. Kang, W. Chang, J. Yoo, and J. C. Ye, "Deep Convolutional Framelet Denosing for Low-Dose CT via Wavelet Residual Network," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1358–1369, June 2018. 5

[18] Dufan Wu, Kyungsang Kim, Georges El Fakhri, and Quanzheng Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2479–2486, 2017. 5

[19] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler, "PWLS-ULTRA: An Efficient Clustering and Learning-Based Approach for Low-Dose 3D CT Image Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1498–1510, June 2018. 5

[20] J. He, Y. Yang, Y. Wang, D. Zeng, Z. Bian, H. Zhang, J. Sun, Z. Xu, and J. Ma, "Optimizing a Parameterized Plug-and-Play ADMM for Iterative Low-Dose CT Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 371–382, Feb 2019. 5

[21] Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, and Daniel Rueckert, "A deep cascade of convolutional neural networks for dynamic mr image reconstruction," *IEEE transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, 2017. 5

[22] Hemant K Aggarwal, Merry P Mani, and Mathews Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 394–405, 2018. 5

[23] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic resonance in medicine*, vol. 79, no. 6, pp. 3055–3071, 2018. 5

[24] Chenyu You, Qingsong Yang, Lars Gjesteby, Guang Li, Shenghong Ju, Zhuiyang Zhang, Zhen Zhao, Yi Zhang, Wenxiang Cong, Ge Wang, et al., "Structurally-sensitive multi-scale deep neural network for low-dose CT denoising," *IEEE Access*, vol. 6, pp. 41839–41855, 2018. 5

[25] Chenyu You, Guang Li, Yi Zhang, Xiaoliu Zhang, Hongming Shan, Mengzhou Li, Shenghong Ju, Zhen Zhao, Zhuiyang Zhang, Wenxiang Cong, et al., "CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)," *IEEE Transactions on Medical Imaging*, vol. 39, no. 1, pp. 188–203, 2019. 5

[26] Chenyu You, Linfeng Yang, Yi Zhang, and Ge Wang, "Low-dose CT via deep CNN with skip connection and network-in-network," in *Developments in X-Ray Tomography XII*. International Society for Optics and Photonics, 2019, vol. 11113, p. 111131W. 5

[27] Qing Lyu, Chenyu You, Hongming Shan, Yi Zhang, and Ge Wang, "Super-resolution MRI and CT through GAN-circle," in *Developments in X-Ray Tomography XII*. International Society for Optics and Photonics, 2019, vol. 11113, p. 111130X. 5

[28] Nimu Yuan, Jian Zhou, Kuang Gong, and Jinyi Qi, "Low-dose CT count-domain denoising via convolutional neural network with filter loss," in *SPIE Medical Imaging: Physics of Medical Imaging*, 2019, vol. 10948, p. 109480R. 5

[29] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, "Deep Learning Computed Tomography: Learning Projection-Domain Weights From Image Domain in Limited Angle Problems," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1454–1463, June 2018. 5

[30] H. Lee, J. Lee, H. Kim, B. Cho, and S. Cho, "Deep-Neural-Network-Based Sinogram Synthesis for Sparse-View CT Image Reconstruction," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 109–119, March 2019. 5

[31] X. Yin, Q. Zhao, J. Liu, W. Yang, J. Yang, G. Quan, Y. Chen, H. Shu, L. Luo, and J. Coatrieux, "Domain Progressive 3D Residual Convolution Network to Improve Low-Dose CT Imaging," *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2903–2913, Dec 2019. 5

[32] Yinsheng Li, Ke Li, Chengzhu Zhang, Juan Montoya, and Guang-Hong Chen, "Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2469–2481, 2019. 5

[33] Huidong Xie, Hongming Shan, Wenxiang Cong, Xiaohua Zhang, Shaohua Liu, Ruola Ning, and Ge Wang, "Dual network architecture for few-view CT-trained on ImageNet data and transferred for medical imaging," in *SPIE Developments in X-Ray Tomography XII*, 2019, vol. 11113, p. 111130V. 5

[34] Cynthia H McCollough, Adam C Bartley, Rickey E Carter, Baiyu Chen, Tammy A Drees, Phillip Edwards, David R Holmes III, Alice E Huang, Farhana Khan, Shuai Leng, et al., "Low-dose CT for the detection and classification of metastatic liver lesions: Results of the

2016 Low Dose CT Grand Challenge," *Medical physics*, vol. 44, no. 10, pp. e339–e352, 2017. 5, 6, 17, 18

[35] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky, "Deep image prior," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 9446–9454. 5, 6, 17, 18, 22, 29

[36] K. Gong, C. Catana, J. Qi, and Q. Li, "PET Image Reconstruction Using Deep Image Prior," *IEEE Transactions on Medical Imaging*, vol. 38, no. 7, pp. 1655–1665, July 2019. 5

[37] Gary Mataev, Michael Elad, and Peyman Milanfar, "DeepRED: Deep Image Prior Powered by RED," in *IEEE International Conference on Computer Vision Workshops(ICCVW)*, 2019. 6

[38] Reinhard Heckel and Paul Hand, "Deep decoder: Concise image representations from untrained non-convolutional networks," *International Conference on Learning Representations*, 2019. 6, 17, 18, 22, 29

[39] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila, "Noise2Noise: Learning Image Restoration without Clean Data," in *International Conference on Machine Learning (ICML)*, 2018, pp. 2965–2974. 6, 17

[40] Shakarim Soltanayev and Se Young Chun, "Training deep learning based denoisers without ground truth data," in *Advances in Neural Information Processing Systems 31*, pp. 3261–3271. 2018. 6, 10, 11

[41] Magauiya Zhussip, Shakarim Soltanayev, and Se Young Chun, "Extending Stein's unbiased risk estimator to train deep denoisers with correlated pairs of noisy images," in *Advances in Neural Information Processing Systems 32*, pp. 1463–1473. 2019. 6

[42] Magauiya Zhussip, Shakarim Soltanayev, and Se Young Chun, "Training deep learning based image denoisers from undersampled measurements without ground truth and without image prior," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 10247–56. 6

[43] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007. 6, 17, 18, 21, 22, 29

[44] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug, "Noise2void-learning denoising from single noisy images," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2129–2137. 6

[45] Joshua Batson and Loic Royer, "Noise2Self: Blind denoising by self-supervision," in *International Conference on Machine Learning (ICML)*, 2019, pp. 524–533. 6

[46] Samuli Laine, Jaakko Lehtinen, and Timo Aila, "Improved self-supervised deep image denoising," in *International Conference on Learning Representations (ICLR) Workshops*, 2019. 6

[47] Yoann Le Montagner, Elsa D Angelini, and Jean-Christophe Olivo-Marin, "An unbiased risk estimator for image denoising in the presence of mixed poisson–gaussian noise," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1255–1268, 2014. 6, 13

[48] J. A. Fessler, "Hybrid Poisson/polynomial objective functions for tomographic image reconstruction from transmission scans," *IEEE Transactions on Image Processing*, vol. 4, no. 10, pp. 1439–1450, Oct 1995. 6, 15, 16

[49] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, May 2017. 7, 18

[50] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T Barron, "Unprocessing images for learned raw denoising," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 11036–11045. 7, 17

[51] Léon Bottou, "Online Learning and Stochastic Approximations," in *On-line learning in neural networks*, pp. 9–42. Cambridge University Press New York, NY, USA, 1998. 8

[52] Yurii Nesterov, "A method of solving a convex programming problem with convergence rate o (1/k2)," in *Soviet Mathematics Doklady*, 1983. 8

[53] Diederik P Kingma and Jimmy Ba, "Adam - A Method for Stochastic Optimization," in *International Conference on Learning Representation (ICLR)*, 2015. 8

[54] C M Stein, "Estimation of the mean of a multivariate normal distribution," *The Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, Nov. 1981. 8

[55] T Blu and F Luisier, "The SURE-LET Approach to Image Denoising," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2778–2786, Oct. 2007. 8

[56] S Ramani, T Blu, and M Unser, "Monte-Carlo Sure: A Black-Box Optimization of Regularization Parameters for General Denoising Algorithms," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1540–1554, Aug. 2008. 9

[57] Markku Makitalo and Alessandro Foi, "Optimal inversion of the anscombe transformation in low-count poisson image denoising," *IEEE transactions on Image Processing*, vol. 20, no. 1, pp. 99–109, 2011. 20, 29

[58] Y C Eldar, "Generalized SURE for Exponential Families: Applications to Regularization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 471–481, Jan. 2009. 31

[59] Martin Raphan and Eero P Simoncelli, "Learning to be Bayesian without supervision," in *Advances in Neural Information Processing Systems (NIPS)*, 2007, pp. 1145–1152. 31

Acknowledgement

저의 짧으면서도 길었던 석사생활을 마치며 제 주변 모든사람들에게 감사인사를 전하고 싶습니다. 먼저 제가 인생에서 앞으로 나아가야 할 길에서 헤매고 있을 때, 고민을 들어주시고 공부할 수 있는 기회를 주신 전세영 교수님께 감사드립니다. 석사생활을 이렇게 잘 마무리 할 수 있었던 것도, 제가 하고싶은 연구를 주도적으로 할 수 있었던 것도 모두 다 교수님 덕분이었고, 단순히 학생이 아니라 한명의 인간으로서 항상 존중 해주심에 감사드립니다. 앞으로 제가 살아가는 데 있어서 저 또한 타인에게 그러한 인물이 될 수 있도록 노력하겠습니다.

그리고 석사생활 동안 저와 가장 많은 시간을 공유하고 이제는 가족같은 BMIPL 식구들에게도 감사를 전하고 싶습니다. 맏형으로서 듬직한 한빛 형, 석사 생활 시작 전부터 연구적으로도, 정신적으로도 항상 의지가 되는 동원 형, 항상 배려해주고 챙겨줬던 지수 형, 묵묵히 이야기 잘들어주고 진국인 원재, 오랫동안 항상 편하고 변함없는게 신기한 동운, 매번 이야기하는 게 즐겁고 공감해주는 용혁, 랩실에서 항상 열심히 하는 병현, 지예 에게 고맙다는 말을 하고싶습니다.

그리고 오랜만에 연락을 해도 항상 재밌게 대화 할 수 있는 상언, 학부생 때부터 지금까지도 정말 좋은 친구들이고 오랫동안 보고싶은 우년, 재은, 정훈, 예영에게도 감사하다는 말을 전하고 싶습니다. 또한 대학원 생활동안 정말 건강하게, 즐겁게 운동할 수있게 도와준 상기 형, 그리고 체육관 식구들에게도 감사합니다.

제가 가장 사랑하는 저의 가족 아빠, 엄마, 형에게도 감사인사를 전하고 싶습니다. 매번 힘들 때, 즐거울 때 저의 이야기를 잘 들어주시고 조언을 해주시면서 저에게 가장 큰 힘이 되는 부모님, 저에게는 이제 가장 가까운 친구이자, 자신의 길을 정말 잘 가고 있어서 고맙고 듬직한 형에게 정말 고맙다는 말을 전하고 싶습니다.

마지막으로, 저의 가장 가까이에서 모든 감정, 모든 일들을 공감해주고, 들어주며 이루 말할 수 없을 만큼 의지가 되고 힘이 되어주는 저의 여자친구 슬기에게 너무나도 고맙습니다. 가장 기쁠 때나, 힘이 든 일이 있을 때 누구보다 먼저 저의 이야기를 들어주고, 응원해주었던 여자친구 덕분에

석사생활을 잘 마무리 하고 졸업할 수 있었습니다. 저의 주변 모든 분들에게 감사하고 앞으로도 더욱 발전하는 제가 되도록 하겠습니다.